Identifying and Estimating Models of Optimal Constracting: Part 2

Robert A. Miller

AES Summer School 2020

July 2020

Motivation

Background

- SOX was a legislative response taken in 2002 by the U.S. government and applied to almost all public firms meant to improve:
    - disclosure quality
    - corporate governance
- Compensation is a crucial mechanism in corporate governance and thus an important angle in any overall evaluation of SOX.
- Through which channels did SOX affect CEO compensation?
    - This paper explicitly attributes SOX effect to changes in different types of agency cost and factors driven by changes in primitives.
    - Measures are derived from a dynamic principal-agent model of moral hazard and hidden information.
- To what extent did SOX affect CEO compensation?
    - This paper quantifies and compares the contribution of each channel

▸ literature

- Main sources: ExecuComp, CRSP and Compustat.
- Sample: S&P1500 firms, 1993 to 2005.
- 12 firm types (denoted as Z_{nt} for firm n in period t):
    - three sectors (primary, consumer goods and services) based on GICS code
    - 2×2 categories in each sector defined by size (A: total assets) and capital structure (C: debt/equity) with indicators L and S

firm characteristics

- Accounting return acc_ret_{nt} for firm n in period t:

$$acc\_ret_{nt} = \frac{Assets_{nt} - Debt_{nt} + Dividend_{nt}}{Assets_{n,t-1} - Debt_{n,t-1}}$$

- Proxy of private state $s_{nt} \in \{1, 2\}$ (CEOs observe and report):

$$s_{nt} \equiv \begin{cases} 1 \text{ (bad)} & \text{if } acc\_ret_{nt} < mean(acc\_ret \mid Z) \\ 2 \text{ (good)} & \text{otherwise} \end{cases}$$

- Performance measure:
  - gross abnormal return $x_{nt} \equiv \widetilde{x}_{nt} + w_{nt} / V_{n,t-1}$, with probability distribution estimated using kernel density.
- Total compensation ($\widetilde{w}_{nt}$):
  - change in wealth (Antle and Smith 1985, 1986, Hall and Liebman 1998, Margiotta and Miller 2000, etc.).
- Optimal compensation: a function of gross abnormal return conditional on bond price, public and private states

$$w_t(x_n | Z, S, b_t) = \frac{\sum_{m=1}^{N} \widetilde{w}_{mt} I\{Z_m = Z, S_m = S\} K\left(\frac{x_{mt} - x_{nt}}{h_x}\right) K\left(\frac{b_m - b_t}{h_b}\right)}{\sum_{m=1}^{N} I\{Z_m = Z, S_m = S\} K\left(\frac{x_{mt} - x_{nt}}{h_x}\right) K\left(\frac{b_m - b_t}{h_b}\right)}$$

  - b_t (bond price) $=$ the present value of an annuity of \$1 Treasury Bill paid for 30 years.

Testing for Structural Change

Compensation (Table 1)

- Accounting matters: compensation in bad states is lower than in good states.
- Primary sector: CEO compensation significantly increased after SOX.
- Sox compressed compensation is all sectors

		Bad			Good		
Sector	(A,C)	Pre	Post	t/F	Pre	Post	t/F
Primary	(S,S)	521	1526	1.3	4096	7206	2.7
		(8244)	(12516)	0.4	(15743)	(14913)	1.1
	(S,L)	-29	2432	4.0	3453	6299	2.3
		(7356)	(5528)	1.8	(9835)	(11433)	0.7
	(L,S)	3429	6849	2.7	6501	9607	2.0
		(10015)	(14023)	0.5	(14415)	(16390)	0.8
	(L,L)	3368	6586	3.9	6015	10569	4.1
		(10847)	(13337)	0.7	(12867)	(17228)	0.6
Consumer Goods	(S,S)	-1203	-1594	-0.4	7301	7494	0.1
		(15811)	(11796)	1.8	(27128)	(22425)	1.5
	(S,L)	-567	65	0.5	4494	4528	0.0
		(11016)	(8268)	1.8	(16528)	(13700)	1.5
	(L,S)	1569	2279	0.3	12432	13560	0.3
		(23109)	(26279)	0.8	(38644)	(33589)	1.3
	(L,L)	3858	5588	0.9	12923	16759	1.4
		(21787)	(25096)	0.8	(31014)	(35861)	0.7
Service	(S,S)	454	558	0.1	6929	7782	0.7
		(14951)	(13127)	1.3	(25125)	(23095)	1.2
	(S,L)	1814	2600	0.6	5145	6497	0.5
		(13211)	(12876)	1.1	(19218)	(23743)	0.7
	(L,S)	5351	5502	0.1	18610	17536	-0.3
		(30923)	(29914)	1.1	(46350)	(37877)	1.5
	(L,L)	6666	6964	0.3	14321	16451	1.1
		(22752)	(19732)	1.3	(29764)	(33072)	0.8

Testing for Structural Change

Two Types of Change

- A structural change in compensation occurs if
  - the distribution of abnormal returns changes
  - the relationship between abnormal returns and CEO compensation changes
- We need to test for equality, between the pre-and post-SOX eras, of
  - the probability density functions for gross abnormal returns
  - the shape of the compensation schedule

- Denote the set of 24 categorical variables (formed from 3 sectors, 2 firm sizes, 2 capital structures, and 2 accounting states) by Z,
- Let $f_{pre}(x_{nt}|z_{nt})$ denote the probability density function of abnormal returns in the pre-Sox era conditional on $z_{nt} \in Z$.
- Also define $f_{post}(x_{nt}|z_{nt})$ in a similar manner.
- Under the null hypothesis of no change $f_{pre}(x|z) = f_{post}(x|z)$ for all $(x, z) \in \mathcal{R} \times Z$.
- Li and Racine (2007, page 363) propose a one-sided test for the null, in which the test statistic is asymptotically distributed standard normal.

nonp results

Testing for Structural Change

Change in the shape of the contract:

- Let $w_{pre}(x_{nt}, z_{nt})$ denote CEO compensation as a function of (x_{nt}, z_{nt}) in the pre-SOX era.
- Similarly define $w_{post}(x_{nt}, z_{nt})$ in the post-SOX era.
- We can test whether the two mappings are equal by including an indicator variable for the post-SOX regime in nonparametric regressions of compensation on the gross abnormal return x_{nt} for each z_{nt}.
- The one-sided test of the null hypothesis of equality is asymptotically standard normal.

▸ nonp results

Testing for Structural Change

Nonparametric Tests Results (Table 2)

A: Test on PDF of Gross Abnormal Returns						
Sector	Primary		consumer		Service	
(A,C)	Bad	Good	Bad	Good	Bad	Good
(S,S)	18.05	10.34	12.51	12.39	14.25	14.55
(S,L)	5.88	5.02	1.26	2.27	14.70	5.29
(L,S)	3.29	4.16	3.74	2.03	9.01	19.69
(L,L)	29.46	8.57	9.03	8.68	71.68	29.56

backnonp result

B: Test on Contract Shape						
Sector	Primary		Consumer		Service	
Firm Type	Bad	Good	Bad	Good	Bad	Good
(S,S)	10.06	1.58	2.89	1.09	1.54	1.47
(S,L)	6.82	6.45	3.30	1.71	4.08	6.85
(L,S)	19.67	7.34	5.51	3.52	5.66	8.74
(L,L)	10.32	23.38	3.69	6.74	7.37	10.65

back1nonp result

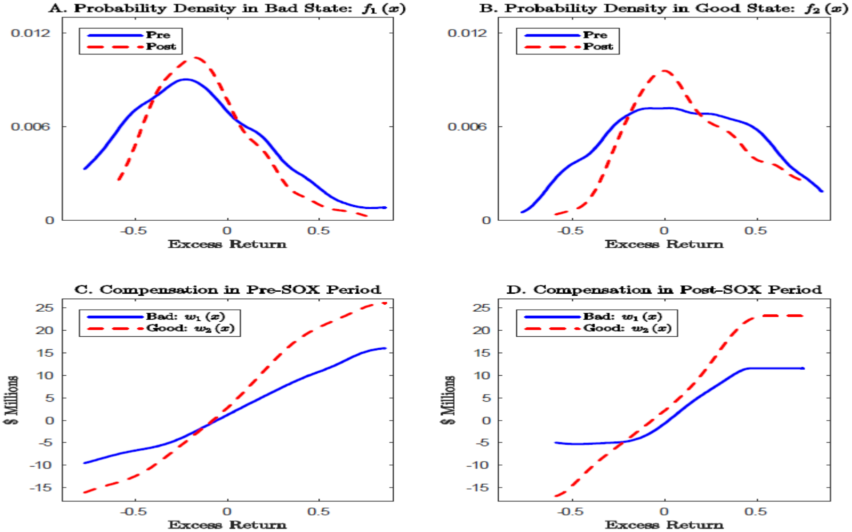
Note: Statistics in both one-sided tests follow N(0,1).

The critical value of 5% confidence level is 1.64.

- Changes occur in both tests in almost all cases among the 3 sectors.

Testing for Structural Change

Nonparametric density and compensation schedules



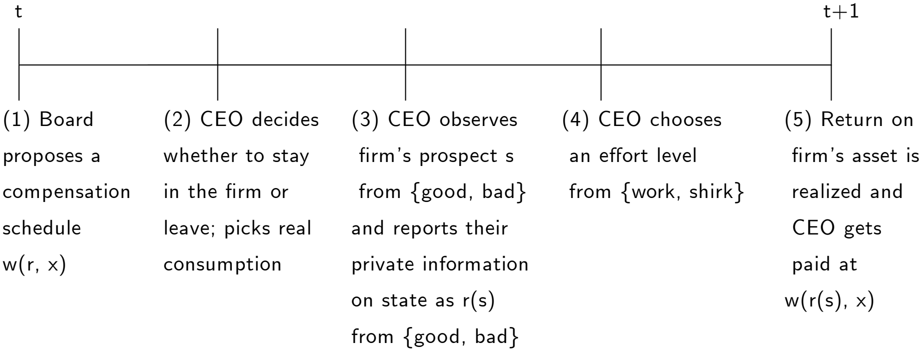
Testing for Structural Change

Where we need a model

- Was the reaction in compensation design a response to
  - the new distribution of abnormal returns?
  - CEO function changes?
- We estimate a dynamic principal-agent model of CEO compensation which is built on primitives that might change with the implementation of SOX.
- It has four key features:
  1. Moral hazard is one contracting friction (Hidden actions are primal in explaining equity-based compensation.)
  2. Hidden information is an additional contracting friction (CEOs benefit from hidden information, Gayle and Miller 2015).
  3. Accounting is a channel for CEOs to reveal private information.
  4. Long-term contract is decentralized to a series of short-term contracts

Model

Timeline



Model

CEO's Indirect Utility in a one period model

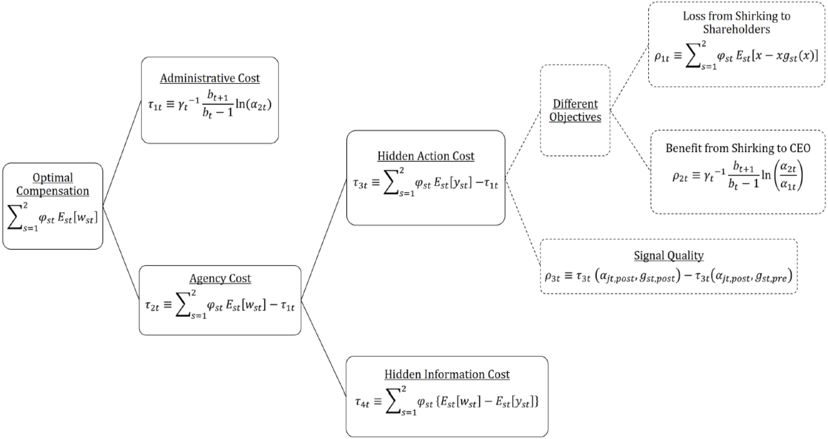
- In a one period model the CEO optimally chooses from {reject, shirk, work} to maximize

$$E[U \mid \text{action, report}] \equiv \begin{cases} -1 & \text{if } \textit{reject} \\ -\sum_{s=1}^{2} \varphi_{st} \left\{ \alpha_{1t}^{\frac{1}{b_t - 1}} \int_{\underline{x}}^{\infty} \exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) g_{st}(x) f_{st}(x) dx \right\} & \text{if } \textit{shirk} \\ -\sum_{s=1}^{2} \varphi_{st} \left\{ \alpha_{2t}{}^{\frac{1}{b_t - 1}} \int_{\underline{x}}^{\infty} \exp\left(-\frac{\gamma_t w_{r_t(s)t}(x)}{b_{t+1}} \right) f_{st}(x) dx \right\} & \text{if } \textit{work} \end{cases}$$

  - φ_{st}: the probability of privately observed state $s_t \in \{1,2\}$
  - γ_t: the coefficient of absolute risk aversion
  - α_{jt}: effort cost parameter for effort choice $j \in \{1,2\}$
  - $f_{st}(x)$: density of performance measure (x) on the equilibrium path
  - $g_{st}(x)$: likelihood ratio of the density of x off the equilibrium path over that on the equilibrium path tech assumptions
  - b_t: the price of a bond that pays a unit of consumption each period from period t onwards, relative to the price of a unit of consumption in period t.

Measuring SOX Effects

Framework



τ_{1t} : optimal compensation if there is neither hidden action nor hidden information.

$y_{st}(x)$: optimal compensation if there is hidden action only.

Optimal Contract

Shareholders solve two problems

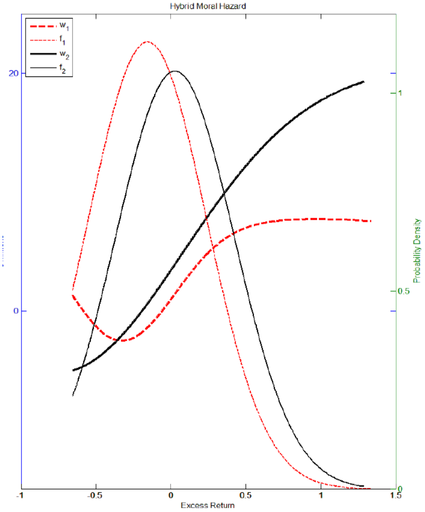
- Minimize expected compensation subject to
  - participation constraint (in both states):
    $E[U \mid \text{work, honest}] - \text{ outside option} \geq 0$
  - incentive compatibility constraint (in both states):
    $E[U \mid \text{work, honest}] - E[U \mid \text{shirk, honest}] \geq 0$
  - truth-telling constraint (in good state):
    $E[U \mid \text{work, honest}] - E[U \mid \text{work, under-report}] \geq 0$
  - sincerity constraint (in good state):
    $E[U \mid \text{work, honest}] - E[U \mid \text{shirk, under-report}] \geq 0$
- Maximize expected net benefits after compensation given CEO works in both private states among all possible effort choices of the CEO.

formal   optimal contract

back1constraints

Optimal Contract

The Optimal Compensation Schedules



Note: The excess return is approximated by one-side truncated normal distribution .

- From the data:
  - we have known the joint distribution of performance measure and compensation
- From the model:
  - we have developed several measures relevant to evaluating SOX, which can be quantified from the data if parameter values are known
- Before estimation, we need to answer: does the model have empirical content?
  - *model specification test*: the model can be rejected if there exist no parameter values that satisfy the restrictions implied by the model.
  - *criterion function*: summarize model restrictions and penalize deviations

- Step 1: Shrink the parameter space
  - equality restrictions involving both observables and unobserved parameters $\Rightarrow$ a platform to derive the mappings from the risk aversion parameter γ to other primitives α_{1t}, α_{2t}, $g_{1t}(x)$, $g_{2t}(x)$
- Step 2: Set identify γ
  - equality and inequality restrictions bound γ up to a set rather than a point
  - construct a criterion function $Q(\gamma)$ to summarize these restrictions

- Unknown primitives: γ, α_{1t}, α_{2t}, $g_{1t}(x)$, $g_{2t}(x)$
- Equality restrictions:
  - First order conditions
  - Binding constraints: Participation, Incentive Compatibility
  - Complementary-slackness conditions: Truth-telling, Sincerity
- Mappings: $v_{st}(x,\gamma) \equiv \exp(-\gamma w_{st}(x)/b_{t+1})$ and $\overline{v}_{st}(\gamma) \equiv \exp(-\gamma \overline{w}_{st}/b_{t+1})$

$$\begin{aligned}
\alpha_{1t} &= \alpha_{2t} \left[\frac{\overline{v}_{2t}(\gamma)^{-1} - E_2\left[v_{2t}(x,\gamma)^{-1}\right]}{\overline{v}_{2t}(\gamma)^{-1} - E_2\left[v_{2t}(x,\gamma)\right]^{-1}} \right]^{b_t - 1} \\
\alpha_{2t} &= E\left[v_{st}(x,\gamma)\right]^{1-b_t} \\
g_{1t}(x) &= \frac{\left\{ \overline{v}_{1t}(\gamma)^{-1} - v_{1t}(x,\gamma)^{-1} + \eta_{3t}\left[\overline{h}_t - h_t(x)\right] \right\} \left[\frac{\alpha_{1t}}{\alpha_{2t}}\right]^{\frac{1}{1-b_t}} - \eta_{4t} g_{2t}(x) h_t(x)}{\overline{v}_{1t}(\gamma)^{-1} - E\left[v_{st}(x,\gamma)\right]^{-1} + \eta_{3t}\overline{h}_t} \\
g_{2t}(x) &= \frac{\overline{v}_{2t}(\gamma)^{-1} - v_{2t}(x,\gamma)^{-1}}{\overline{v}_{2t}(\gamma)^{-1} - E_2\left[v_{2t}(x,\gamma)^{-1}\right]}
\end{aligned}$$

identification of KT multipliers | identification illustration | optimal contract

Identification

Step 2: Set Identification

- Equality restrictions: $\Psi_{lt}(\gamma^*)=0$
  - Likelihood ratio is non-negative with unit mass: $\Psi_{1t}(\gamma^*)=E_1[I\{g_{1t}(x,\gamma^*)\geq 0\}-1]=0$
  - α_{1t} is the same between two private states
  - Between truth-telling and sincerity, there is at least one binding
  - Complementary-slackness conditions for truth-telling and sincerity
- Inequality restrictions: $\Lambda_{kt}(\gamma^*)\geq 0$
  - Shareholders prefer working to shirking: $\Lambda_{2t}(\gamma^*)=E[(Vx-w)|working]-E[(Vx-w)|shirking]$
  - Kuhn Tucker multipliers are positive
- Identified set:

$$\Gamma \equiv \left\{ \gamma > 0 : Q(\gamma) \equiv \sum_{t=1}^{T} \sum_{l} \Psi_{lt}(\gamma)^2 + \sum_{t=1}^{T} \sum_{k} \min\left[0, \Lambda_{kt}(\gamma)\right]^2 = 0 \right\}$$

▸ constraints ▸ identification of KT multipliers

- Determined by the critical value c_δ associated with test size δ:

$$\Gamma_\delta^{(N)} \equiv \left\{ \gamma > 0 : Q^{(N)}(\gamma) \leq c_\delta \right\}$$

  - Intuition: an estimate of γ is penalized if it makes the sample analogue of the criterion function deviate from its theoretical value (zero).
  - How: a consistent estimate of c_δ can be determined numerically by following subsampling procedures described in the paper, which follows Chernozhukov Hong and Tamer (2007).

Results

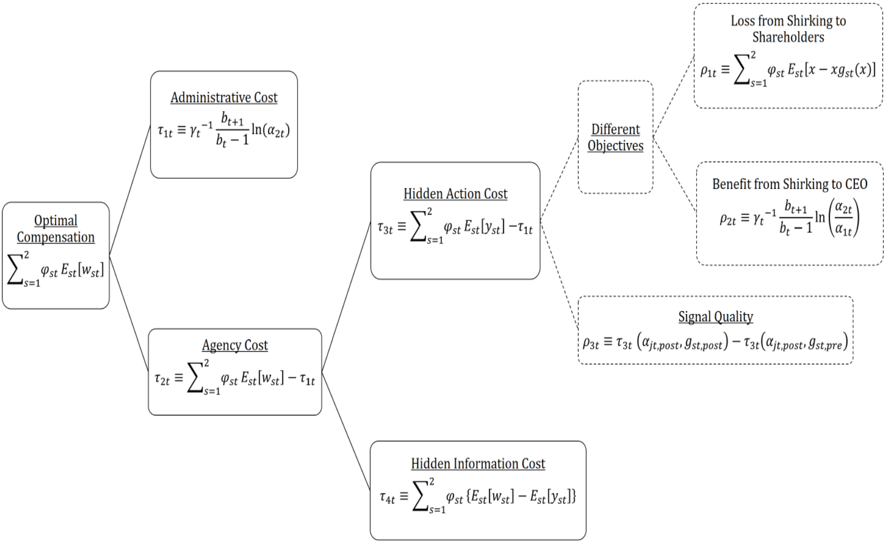
The 95% Confidence Regions of Risk-aversion

Period	Years	Risk Aversion	Certainty Equivalent
Pre	1993-2001	(0.0695, 0.6158)	(34722, 290206)
Post	2004-2005	(0.0695, 0.6158)	(34722, 290206)

- The estimated risk aversion is comparable to previous studies in terms of certainty equivalent which is the dollar amount that a CEO at that risk aversion level would like to pay to avoid a gamble that has half chance of losing or winning $1 million.
- We do not reject the null hypothesis that no significant change in risk attitude occurred after SOX.

Results

Framework



Results

Administrative Cost (measured in thousands of 2006 US$, Table 3)

Sector	(A,C)	$\Delta\tau_1$	
Primary	(S,S)	(2076, 2194)	+
	(S,L)	(3692, 3879)	+
	(L,S)	(6254, 7054)	+
	(L,L)	(1846, 2181)	+
Consumer	(S,S)	(2203, 2438)	+
	(S,L)	(587, 614)	+
	(L,S)	(-3929, -2025)	-
	(L,L)	(218, 581)	+
Service	(S,S)	(-1819, -1295)	-
	(S,L)	(-230, 158)	=
	(L,S)	(-5403, -4827)	-
	(L,L)	(-1742, -1191)	-

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\tau_1 \equiv \gamma^{-1} \frac{b_{t+1}}{b_t - 1} \ln \alpha_{2,pre}$.

▸ framework

Results

Aggregate Agency Costs (measured in thousands of 2006 US$, Table 4)

Sector	(A,C)	$\Delta\tau_2$	
Primary	(S,S)	(20, 190)	+
	(S,L)	(3, 30)	+
	(L,S)	(76, 611)	+
	(L,L)	(43, 379)	+
Consumer	(S,S)	(-527, -59)	-
	(S,L)	(21, 156)	+
	(L,S)	(182, 1812)	+
	(L,L)	(81, 459)	+
Service	(S,S)	(-360, -41)	-
	(S,L)	(45, 395)	+
	(L,S)	(113, 355)	+
	(L,L)	(53, 529)	+

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\tau_2 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}\left[w_{s,pre}(x)\right] - \tau_1.$

framework

Results

Welfare Costs of Moral Hazard (measured in thousands of 2006 US$, Table 5)

Sector	(A,C)	$\Delta\tau_3$	
Primary	(S,S)	(228, 1532)	+
	(S,L)	(-39, 165)	=
	(L,S)	(96, 1774)	+
	(L,L)	(265, 380)	+
Consumer	(S,S)	(-4387, -600)	-
	(S,L)	(-817, -202)	-
	(L,S)	(-3111, -649)	-
	(L,L)	(-3848, -332)	-
Service	(S,S)	(-328, -150)	-
	(S,L)	(-399, 268)	=
	(L,S)	(-6438, 433)	=
	(L,L)	(-2621, 479)	=

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\tau_3 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}\left[y_{s,pre}(x)\right] - \tau_1.$

framework

Results

Welfare Costs of Hidden Information (in thousands of 2006 US$, Table 6)

Sector	(A,C)	$\Delta\tau_4$	
Primary	(S,S)	(-1342, -208)	-
	(S,L)	(-135, 42)	=
	(L,S)	(-1163, -20)	-
	(L,L)	(-227, 87)	=
Consumer	(S,S)	(540, 3860)	+
	(S,L)	(223, 973)	+
	(L,S)	(831, 4923)	+
	(L,L)	(413, 4307)	+
Service	(S,S)	(-32, 217)	=
	(S,L)	(-218, 795)	=
	(L,S)	(-320, 6788)	=
	(L,L)	(-348, 3150)	=

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\tau_4 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre} \left[w_{s,pre}(x) - y_{s,pre}(x) \right]$.

framework

Results

Shareholders' Losses from Shirking (%, Table 7)

Sector	(A,C)	$\Delta\rho_1$	
Primary	(S,S)	(-2.69, -1.96)	-
	(S,L)	(-6.92, -4.75)	-
	(L,S)	(-2.82, -2.10)	-
	(L,L)	(-1.96, -1.95)	-
Consumer	(S,S)	(-9.16, -8.72)	-
	(S,L)	(2.12, 12.21)	+
	(L,S)	(-0.40, 1.54)	-
	(L,L)	(-2.68, -2.11)	-
Service	(S,S)	(-8.93, -6.34)	-
	(S,L)	(-3.02, -1.03)	-
	(L,S)	(-16.59, -15.37)	-
	(L,L)	(-5.97, -5.07)	-

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\rho_1 \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre} \{x [1 - g_{s,pre}(x)]\}$.

▸ framework

Results

CEOs' Benefit from Shirking (in thousands of 2006 US$, Table 8)

Sector	(A,C)	$\Delta\rho_2$	
Primary	(S,S)	(122, 221)	+
	(S,L)	(-57, -24)	-
	(L,S)	(1716, 2125)	+
	(L,L)	(100, 380)	+
Consumer	(S,S)	(-3213, -2091)	-
	(S,L)	(287, 476)	+
	(L,S)	(18, 792)	+
	(L,L)	(-1078, -654)	-
Service	(S,S)	(-780, -487)	-
	(S,L)	(67, 446)	+
	(L,S)	(-7697, -5721)	-
	(L,L)	(-2041, -1985)	-

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\rho_2 \equiv b_{t+1} \left[(b_t - 1) \gamma \right]^{-1} \ln(\alpha_{2,pre} / \alpha_{1,pre})$.

framework

Results

Change in Moral Hazard Cost due to Signal Quality (in thousands of 2006 US$, Table 9)

Sector	(A,C)	ρ_3	
Primary	(S,S)	(216, 1261)	+
	(S,L)	(-37, 198)	=
	(L,S)	(-617, -151)	-
	(L,L)	(59, 225)	+
Consumer	(S,S)	(180, 1254)	+
	(S,L)	(-1575, -272)	-
	(L,S)	(-3202, -930)	-
	(L,L)	(-224, 1181)	=
Service	(S,S)	(-130, 398)	=
	(S,L)	(-503, 212)	=
	(L,S)	(1503, 24282)	+
	(L,L)	(724, 4470)	+

Note: A: Size. C: Debt/Equity. S: Small. L: Large.

$\rho_3 \equiv \tau_3(\alpha_{j,post}, g_{s,post}) - \tau_3(\alpha_{j,post}, g_{s,pre})$.

$\tau_3(\alpha_{j,post}, g_{s,pre}) \equiv \sum_{s=1}^{2} \varphi_{s,pre} E_{s,pre}[y_s(x, \alpha_{j,post}, g_{s,pre})] - \tau_1(\alpha_{j,post}, g_{s,pre})$.

framework optimal contract formula

- The structural changes in CEO compensation across 2002 could be attributed to
  (1) change in CEOs' risk attitude
  (2) change in aggregate economy
- We rule out (1) by showing no significant change in the estimated risk aversion.
  - contradicting concerns that CEOs would overreact to SOX provisions and exercise undue caution in investment decisions thus destroying shareholder value. (Coats and Srinivasan 2014).
- We rule out (2) by anchoring the estimation of agency costs at the same bond prices for both pre and post periods.

Conclusion

Our Explanations

- We attribute the structural changes in CEO compensation to the new regulations passed in 2002.
- The main impact of SOX was to increase the administrative burden from compliance in small firms, and in the primary sector more generally.
  - complementing Coats and Srinivasan (2014) who document the direct costs from control system expenditures incurred due to SOX's new requirements.
- SOX also increased agency costs in large firms, as well as in the primary sector more generally. The most important reason for the increased agency costs in the primary sector proved to be the higher compensating differential between shirking and working.

Conclusion

Policy Implication

- SOX effect is firm type dependent
  - Primary: both administrative and agency costs increased within all firm types
  - Consumer goods and services: for small firms, declines in agency costs roughly offset increases in administrative costs, whereas for large firms, it was the other way around.
- Policy implication: there may be scope for relaxing some of the regulations that SOX introduced, either by rolling back its harsher provisions, or tailoring them to different firm types.

- - - - Backup Slides - - - -

Motivation

Literature on SOX Effect

- firm behavior: extensive evidence
  - switching earnings management methods (Cohen et al 2008), reducing investment (Bargeron et al. 2010, Cohen et al. 2007, Kang et al. 2010), delisting (Engel et al. 2006, Leuz et al. 2007)
- stock market reaction: mixed evidence
  - Zhang 2007, Jain and Rezaee 2006, Leuz 2007, Dey 2010, Livtak 2007, Hochberg et al. 2009
- compensation practice: crucial but underexplored
  - exceptions: Carter et al. 2009, Nekipelov 2010, Cohen et al. 2007

▸ from literature

backfirm characteristics

Table 1: Cross-section Summary of Firm Characteristics

(Asset in millions of 2006 US$, Standard deviations in parentheses)

Sector	Primary		Consumer		Service	
Period	(1) Pre	(2) Post	(3) Pre	(4) Post	(5) Pre	(6) Post
Total Asset	4597	6477	3516	5168	13499	17568
	(7677)	(11117)	(8103)	(12439)	(42502)	(70163)
Debt/Equity	1.819	2.102	1.571	1.537	3.650	2.982
	(1.450)	(2.575)	(1.646)	(2.119)	(5.263)	(4.323)
Accounting Return	1.108	1.138	1.126	1.108	1.175	1.115
	(0.228)	(0.217)	(0.282)	(0.253)	(0.335)	(0.261)
Abnormal Return	-0.031	0.063	-0.029	0.014	0.022	0.044
	(0.319)	(0.277)	(0.383)	(0.320)	(0.433)	(0.353)
Observations	6213	1620	5012	1394	7629	2656

The 95% Confidence Regions of Risk-aversion and Corresponding Certainty Equivalent (in $)

A: Full Sample			
Period	Years	Risk Aversion	Certainty Equivalent
Pre	1993-2002	(0.0784, 0.2335)	(39160, 115704)
Post	2003-2005	(0.0616, 0.2335)	(30781, 115704)
Common		(0.0784, 0.2335)	(39160, 115704)
B: Restricted Sample			
Period	Years	Risk Aversion	Certainty Equivalent
Pre	1993-2001	(0.0695, 0.6158)	(34722, 290206)
Post	2004-2005	(0.0695, 0.6158)	(34722, 290206)
Common		(0.0695, 0.6158)	(34722, 290206)

Note: The subsampling procedure was performed using 100 replications of subsamples with 80 percent of full sample observations, each using 100 grid points on the searching interval [0.0003,54.598]. The certainty equivalent corresponding to one particular value of the risk aversion in the estimated confidence region is the dollar amount that a CEO at that risk aversion level would like to pay to avoid a gamble that has an equal probability of losing or winning one million dollars.

▸ backtech assumptions

- For each private state $s_t \in \{1,2\}$, denote by $f_{st}(x)$ the probability density function for return conditional on the agent working, and let $f_{st}(x) g_{st}(x)$ denote the probability density function for return when the agent shirks. Then $g_{st}(x)$, the ratio of the two densities, is a likelihood ratio and is nonnegative for all x:

$$E_{st}\left[g_{st}(x)\right] \equiv \int g_{st}(x) f_{st}(x)\, dx = 1$$

- We assume that the principal prefers the agent working to shirking

$$E_{st}\left[x g_{st}(x)\right] \equiv \int x f_{st}(x) g_{st}(x)\, dx < \int x f_{st}(x)\, dx \equiv E_{st}[x]$$

- We assume the likelihood of shirking declines to zero as abnormal returns increase without bound:

$$\lim_{x \to \infty} \left[g_{st}(x)\right] = 0$$

- We assume the weighted likelihood ratio of the second state occurring relative to the first given any observed value of excess returns, $x \in R$ converges to an upper finite limit as x increases, such that:

$$\lim_{x \to \infty} \left[\varphi_{2t} f_{2t}(x) / \varphi_{1t} f_{1t}(x)\right] \equiv \lim_{x \to \infty} \left[h_t(x)\right] = \sup_{x \in R} \left[h_t(x)\right] \equiv \overline{h}_t < \infty.$$

- To induce an honest, diligent manager to participate, her expected utility from employment must exceed the utility she would obtain from outside option:

$$\alpha_{2t}^{1/(b_t-1)} \left[\sum_{s=1}^{2} \int_{\underline{x}}^{\infty} \varphi_{st} \exp\left(-\gamma_t w_{st}(x)/b_{t+1}\right) f_{st}(x)\, dx \right] \leq 1$$

- Given her decision to stay with the firm one more period, and to truthfully reveal the state, the incentive compatibility constraint induces the manager to prefer working diligently to shirking in each state $s_t \in \{1, 2\}$:

$$\begin{aligned} & \alpha_{2t}^{1/(b_t-1)} \int_{\underline{x}}^{\infty} \exp\left(-\gamma_t w_{st}(x)/b_{t+1}\right) f_{st}(x)\, dx \\ \leq \quad & \alpha_{1t}^{1/(b_t-1)} \int_{\underline{x}}^{\infty} \exp\left(-\gamma_t w_{st}(x)/b_{t+1}\right) g_{st}(x)\, f_{st}(x)\, dx \end{aligned}$$

Feasible Contracts

Truth Telling and Sincerity

- Incentives must be provided to persuade the manager not to understate them:

$$\alpha_{2t}^{1/(b_t-1)} \int_{\underline{x}}^{\infty} \exp\left(-\gamma_t w_{2t}(x)/b_{t+1}\right) f_{2t}(x)\, dx$$
$$\leq \quad \alpha_{2t}^{1/(b_t-1)} \int_{\underline{x}}^{\infty} \exp\left(-\gamma_t w_{1t}(x)/b_{t+1}\right) f_{2t}(x)\, dx$$

- An optimal contract also induces the manager not to understate and shirk:

$$\alpha_{2t}^{1/(b_t-1)} \int_{\underline{x}}^{\infty} \exp\left(-\gamma_t w_{2t}(x)/b_{t+1}\right) f_{2t}(x)\, dx$$
$$\leq \quad \alpha_{1t}^{1/(b_t-1)} \int_{\underline{x}}^{\infty} \exp\left(-\gamma_t w_{1t}(x)/b_{t+1}\right) g_{2t}(x) f_{2t}(x)\, dx$$

backformal

Optimal Contract

▸ ρ_3

- If there is neither hidden information nor hidden action, the agent gets $\gamma^{-1}\frac{b_{t+1}}{b_t-1}\ln\alpha_{2t}$
- If there is only hidden action, the optimal compensation derived from a pure moral hazard model is ▸ back1optimal contract

$$y_{st}(x) = \gamma^{-1}\frac{b_{t+1}}{b_t - 1}\ln\alpha_2 + \gamma^{-1}b_{t+1}\ln[1 + \eta^p_{st}(\alpha_{2t}/\alpha_{1t})^{1/(b_t-1)} - \eta^p_{st}g_{st}(x)]$$

where η^p_{st} is the unique positive solution to:

$$\int_{\underline{x}}^{\infty}\frac{g_{st}(x) - (\alpha_{2t}/\alpha_{1t})^{1/(b_t-1)}}{1 + \eta^p_{st}(\alpha_{2t}/\alpha_{1t})^{1/(b_t-1)} - \eta^p_{st}g_{st}(x)}f_s(x)dx = 0$$

- Solving the two problems generates a group of equilibrium restrictions and the optimal contacts expressed in FOCs for each state as

$$\begin{aligned}
\exp\left[\gamma_t w_{1t}(x)/b_{t+1}\right] &= \alpha_{2t}^{1/(b_t-1)} + \eta_{1t}\left[(\alpha_{2t}/\alpha_{1t})^{1/(b_t-1)} - g_{1t}(x)\right] \\
&\quad -\eta_{3t}h_t(x) - \eta_{4t}(\alpha_1/\alpha_2)^{1/(b_t-1)}g_{2t}(x)h_t(x) \\
\exp\left[\gamma_t w_{2t}(x)/b_{t+1}\right] &= \alpha_{2t}^{1/(b_t-1)} + \eta_{2t}\left[(\alpha_{2t}/\alpha_{1t})^{1/(b_t-1)} - g_{2t}(x)\right] + \eta_{3t} + \eta_{4t}
\end{aligned}$$

▸ backoptimal contract ▸ back2optimal contract

Identification

Shadow prices of constraints

- Shadow prices of constraints: η_{0t} (Participation), η_{1t} (Incentive Compatibility for $s_t\!=\!1$), η_{2t} (Incentive Compatibility for $s_t\!=\!2$), η_{3t} (Truth-telling), η_{4t} (Sincerity)
- Define $v_{st}(x) \equiv \exp(-\gamma_t w_{st}(x)/b_{t+1})$

$$\begin{aligned}
\eta_{0t}(\gamma) &= \alpha_{2t}(\gamma) \\
\eta_{1t}(\gamma) &= \left[\frac{\alpha_{1t}(\gamma)}{\alpha_{2t}(\gamma)}\right]^{1/(1-b_t)} \left\{\overline{v}_{1t}(\gamma) - E\left[v_{st}(x,\gamma)\right]^{-1} + \eta_{3t}\overline{h}_t\right\} \\
\eta_{2t}(\gamma) &= \overline{v}_{2t}(\gamma)^{-1} - E_2\left[v_{2t}(x,\gamma)^{-1}\right] \\
\eta_{3t}(\gamma) &= E_2\left[v_{2t}(x,\gamma)\right]^{-1} - \eta_{4t} - E\left[v_{st}(x,\gamma)\right]^{-1} \\
\eta_{4t}(\gamma) &= \frac{\frac{E_1[v_{1t}(x,\gamma)]}{E[v_{st}(x,\gamma)]} - E_1\left[v_{1t}(x,\gamma)h_t(x)\right]\left\{E_2\left[v_{2t}(x,\gamma)\right]^{-1} - E\left[v_{st}(x,\gamma)\right]^{-1}\right\} - 1}{\left[\frac{\alpha_{1t}(\gamma)}{\alpha_{2t}(\gamma)}\right]^{1/(1-b_t)} E_1\left[v_{1t}(x,\gamma)g_{2t}(x,\gamma)h_t(x)\right] - E_1\left[v_{1t}(x,\gamma)h_t(x)\right]}
\end{aligned}$$

▸ backidentification of KT multipliers

▸ back1identification of KT multipliers

Identification

Example

- α_{2t} can be identified from binding participation constraint.
- participation constraint:

$$\alpha_{2t}^{1/(b_t-1)} \left[\sum_{s=1}^{2} \int_{\underline{x}}^{\infty} \varphi_{st} \exp\left(-\gamma_t w_{st}(x) / b_{t+1}\right) f_{st}(x)\, dx \right] \leq 1$$

$$\text{i.e. } \alpha_{2t}^{1/(b_t-1)} E\left[v_{st}(x,\gamma)\right] \leq 1$$

- rearrange items when it is binding to get the mapping

$$\alpha_{2t}(\gamma) = E\left[v_{st}(x,\gamma)\right]^{1-b_t}$$

backidentification illustration

Identification

Criterion Function (detailed)

- Equality restrictions: $\Psi_{jt}(\gamma)=0$
  - Likelihood ratio is non-negative with unit mass: $\Psi_{1t}(\gamma)=E_1[I\{g_{1t}(x,\gamma)\geq 0\}-1]$
  - α_{1t} is the same between two states $s_t\in\{1,2\}$:
    $\Psi_{2t}(\gamma)=E_1[g_{1t}(x,\gamma)v_{1t}(x,\gamma)]/E_1[v_{1t}(x,\gamma)]-E_2[g_{2t}(x,\gamma)v_{2t}(x,\gamma)]/E_2[v_{2t}(x,\gamma)]$
  - Between truth-telling and sincerity, there is at least one binding:
    $\Psi_{3t}(\gamma)*\Psi_{4t}(\gamma)=0$, where $\Psi_{3t}(\gamma)=\alpha_{2t}^{1/(b_t-1)}E_2[v_{2t}(x,\gamma)]-\alpha_{2t}^{1/(b_t-1)}E_2[v_{1t}(x,\gamma)]$ and $\Psi_{4t}(\gamma)=\alpha_{2t}^{1/(b_t-1)}E_2[v_{2t}(x,\gamma)]-\alpha_{1t}^{1/(b_t-1)}E_2[v_{1t}(x,\gamma)g_{2t}(x,\gamma)]$
  - Complementary-slackness conditions for truth-telling and sincerity:
    $\Psi_{3t}(\gamma)*\eta_{3t}(\gamma)=0$, $\Psi_{4t}(\gamma)*\eta_{4t}(\gamma)=0$,
- Inequality restrictions: $\Lambda_{kt}(\gamma)\geq 0$
  - Kuhn Tucker multipliers are positive: $\Lambda_{kt}(\gamma)=\eta_{kt}(\gamma)$ for $k=1,3,4$
  - Shareholders prefer working to shirking:
    $\Lambda_{2t}(\gamma)=E[(Vx-w)|working]-E[(Vx-w)|shirking]$

- Administrative cost: CEOs' pecuniary cost of working under perfect monitoring.

$$\tau_{1t} \equiv \gamma^{-1} \frac{b_{t+1}}{b_t - 1} \ln \alpha_2 .$$

- Aggregate agency cost:

$$\tau_{3t} \equiv E\left[w_{st}(x)\right] - \tau_{1t} = \sum_{s=1}^{2} \int_{\underline{x}}^{\infty} \varphi_s(z) \left[w_{st}(x) - \gamma_t^{-1} \ln \alpha_2\right] f_s(x) dx.$$

- Cost to hidden action only:

$$\tau_{3t} \equiv E_{st}[y_{st}(x)] - \tau_{1t}$$

- Cost to hidden information only:

$$\tau_{4t} \equiv E_s\left[w_{st}(x) - y_{st}(x)\right]$$

- Losses of shareholders from CEO shirking

$$\rho_{1t} \equiv E_{st}\left[x - x g_{st}(x)\right]$$

- Interest alignment: CEO's pecuniary benefit from shirking instead of working under perfect monitoring.

$$\rho_{2t} \equiv \gamma_t^{-1} \frac{b_{t+1}}{b_t - 1} \ln\left(\frac{\alpha_2}{\alpha_1}\right).$$

- Signal quality effect:

$$\rho_{3t} \equiv \tau_3(\alpha_{j,post}, g_{s,post}) - \tau_3(\alpha_{j,post}, g_{s,pre})$$

Identification

A quick summary

- Consolidating the restrictions directly applied to the model, we define:

$$\overline{\Gamma} \equiv \left\{ \gamma > 0 : \begin{array}{l} \Lambda_i(\gamma) \geq 0 \text{ for } i \in \{1,2,3\} \\ \eta_j(\gamma) \geq 0 \text{ for } j \in \{1,3,4\} \\ \Psi_1(\gamma) = 0 \text{ and } \Psi_k(\gamma) \geq 0 \text{ for } k \in \{3,4\} \\ \Psi_3(\gamma)\Psi_4(\gamma) = \Psi_3(\gamma)\eta_3(\gamma) = \Psi_4(\gamma)\eta_4(\gamma) = 0 \end{array} \right\}$$

  where:

  - $\Lambda_i(\gamma)$ refers to suboptimal choices by shareholders of shirking and diligence in the two states (that is shirking in one or both).
  - $\eta_j(\gamma)$ are representations of the Kuhn Tucker multipliers (for incentive compatibility in the first state, truth telling and sincerity).
  - $\Psi_k(\gamma)$ refers to incentive compatibility and sincerity conditions, a positivity condition on $g_1(x,\gamma)$, and a condition that the taste for shirking is independent of the state.

Estimation

$$Q_{\mathrm{H}}\left(\widetilde{\gamma}\right) \equiv \sum_{t=1}^{T} \sum_{j=1}^{9} \min\left[0, \eta_j\left(\gamma\right)\right]^2 + \sum_{t=1}^{T} \sum_{j=6}^{7} \left[\Psi_{5t}\left(\widetilde{\gamma}\right) \Psi_{jt}\left(\widetilde{\gamma}\right)\right]^2$$

$$+ \sum_{t=1}^{T} \Psi_{4t}\left(\widetilde{\gamma}\right)^2 + \sum_{t=1}^{T} \left[\Psi_{6t}\left(\widetilde{\gamma}\right) \Psi_{8t}\left(\widetilde{\gamma}\right)\right]^2 + \sum_{t=1}^{T} \sum_{k=1}^{3} \min\left[0, \Lambda_{kt}\left(\widetilde{\gamma}\right)\right]^2 . \quad (1)$$

- The identified set of risk parameters defined Γ has a simple empirical analogue.
- Suppose we have N cross sectional observations on (x_n, w_n) on identical firms and their managers.
- To estimate $Q(\gamma)$, we replace $\overline{w}$ with $\overline{w}^{(N)} \equiv \max\{w_1, \ldots, w_N\}$ and substitute sample moments for their population corresponding expectations

Empirical Implementation

Convergence of the approximation

- Our tests are based on the fact that if $\gamma \in \Gamma$ then sampling error is the only explanation for why $Q^{(N)}(\gamma)$ might be negative.
- Clearly $Q^{(N)}(\gamma)$ converges at the rate of its slowest converging component.
- For simplicity suppose there exists some $\overline{x} < \infty$ such that $g(x) = 0$ for all $x > \overline{x}$.
- In words, there is a revenue threshold that shirking cannot achieve.
- Thus compensation is flat at $\overline{w}$ for all profits levels above $\overline{x}$, and $\overline{w}^{(N)}$ converges to $\overline{w}$ at a faster rate than $\sqrt{N}$.
- Since all the other components of $Q^{(N)}(\gamma)$ are sample moments, we conclude $Q^{(N)}(\gamma)$ converges at rate $\sqrt{N}$.